DOCUMENT RESUME

ED 320 506 HE 023 597

AUTHOR Swanson, Judy H.
TITLE The Effectiveness of Tutorial Strategies: An Experimental Evaluation.
SPONS AGENCY NCR Corp., Atlanta, GA. Human Interface Center.
PUB DATE Apr 90
NOTE 34p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS College Students; Comparative Analysis; Discovery Learning; Higher Education; Individual Instruction; *Learning Strategies; Lecture Method; *Teaching Methods; *Tutoring; Undergraduate Study

ABSTRACT

The effectiveness of three tutorial strategies was evaluated by experimentally manipulating the strategies used in a 2 X 3 factorial design. Two tutors taught 48 college undergraduates a basic optics lesson about "how lenses work," using three instructional methods which varied the amount of tutor control. In the lecture condition, tutors assumed full control; in the discovery condition, students assumed full control; and in the contingent condition, tutor control was adapted to the needs of individual students. Serving as a baseline, eight students learned the same lesson from a text. Generalized regression analysis was used to assess the effects of aptitudes, treatments, and their interactions (ATI). Significant ATI suggested that discovery methods were particularly ineffective with low ability students, whereas less able students benefited from contingent techniques, and lecture techniques produced intermediate results. When correctly applied, contingent instruction produced the highest learning outcomes across the ability range; however, this strategy proved difficult to carry out consistently. Implications for tutor training and the development of "intelligent" tutoring systems are discussed. Contains 21 references. (Author)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED320506

The Effectiveness of Tutorial Strategies: An Experimental Evaluation

Judy H. Swanson[1]
Stanford University

Paper presented at the annual conference of the American Educational Research Association, Boston, MA. April, 1990.

[1]The research reported in this paper was supported by a grant from NCR Corporation, The Human Interface Center, Atlanta, GA

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JUDY H. SWANSON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

HE 023 597

BEST COPY AVAILABLE

Abstract

The effectiveness of three tutorial strategies was evaluated by experimentally manipulating the strategies used in a 2 x 3 factorial design. Two tutors taught 48 college undergraduates a basic optics lesson about "how lenses work," using three instructional methods, which varied the amount of tutor control. In the lecture condition tutors assumed full control; in the discovery condition, students assumed full control; and in the contingent condition, tutor control was adapted to the needs of individual students. Serving as a baseline, eight students learned the same lessson from a text. Generalized regression analysis was used to assess the effects of aptitudes, treatments, and their interactions (ATI). Significant ATI suggested that discovery methods were particularly ineffective with low ability students, whereas less able students benefited from contingent techniques. Lecture techniques produced intermediate results. When correctly applied contingent instruction produced the highest learning outcomes across the ability range, however, this strategy proved difficult to carry out consistently. Implications for tutor training and the development of "intelligent" tutoring systems are discussed.

Recent successes and disappointments in Intelligent Tutoring Systems (ITS) research underscore the need for a deep understanding of instructional processes. As a result, many researchers (Collins & Steven, 1982; Lepper & Chabay, 1988; Littman, Pinto, & Soloway, 1987; McKendree, Reiser, & Anderson, 1984) have realized that there is much to be learned from studying expert human tutors. Surprisingly little is known about the process of tutoring, and yet, there seems to be a tacit assumption that one-to-one tutoring provides the optimal form of teaching and learning. Bloom (1984) for example, has conducted a number of studies in "search of methods of group instruction as effective as one-to-one tutoring." In this series of studies Bloom used tutoring as the "ideal" standard against which the effectiveness of other, presumably less optimal, teaching methods were assessed.

Although a number of studies have demonstrated the positive effects of tutoring on student learning, most of these (reviewed by Cohen, Kulik, & Kulik, 1982) have been evaluation studies of the form: pretest, tutoring interventions, post-test. These studies fail to address the questions, How do students learn from tutors?, and What is it about the tutoring process that makes it an effective mode of instruction? These are important questions for three reasons. First, because tutoring has become such a popular strategy for improving the academic performance of low-achieving students, we need to understand the process in order to exploit its positive aspects and improve the quality of intervention programs. Second, "expert" human tutors can serve as models for the improvement of computerized "intelligent" tutoring systems. And third, important lessons may be learned for all forms of instruction; it may be possible to incorporate some of the critical features of tutoring into classroom instructional practices. The present study begins to address these questions by experimentally manipulating the instructional strategies used by tutors, thereby making it possible to evaluate the effectiveness of various strategies.

Wood, Bruner, & Ross (1976) developed a schema describing the nature of the tutoring process based on interactions between parents and their young children in informal teaching situations. The intervention of a tutor in this process involves what Bruner (1978) calls a "scaffolding" process that enables a child to solve a problem, or achieve a goal that is initially beyond the child's capacity. Scaffolding consists of the tutor controlling those elements of the task that are too difficult for the child, thereby permitting the child to concentrate upon and complete those elements that the child is able to do on his or her own. By keeping the task demands within the child's "zone of proximal development" (Vygotsky, 1978) the tutor not only helps the child complete the task at hand, but also teaches the child new skills and strategies that will eventually enable the child to accomplish similar tasks alone.

Based on a series of studies, Wood (1980) concluded that a tutor's effectiveness was determined by the extent to which the instructional scaffolding provided by the tutor's interventions were contingent upon the interaction of the task demands and the performance of the student. As defined by Wood, contingent instruction requires that tutorial interventions are inversely related to the student's level of competence. The more difficulty a student is having with a task, the more directive the intervention needs to be. The more success the student experiences, the more the tutor should allow (even push) the student to do on his or her own. The critical point is that the tutor's interventions must be contingent upon the interaction of the task demands and the performance of the student. To do this the tutor must be aware of what the student is attempting to achieve, and be able to diagnose success and failure. Assessing while teaching is difficult because the tutor's model of the student can only be an approximation as the tutor's knowledge of the student is always incomplete.

In Wood's (1980) studies, three-to-four year olds were taught how to construct a

three-dimensional puzzle. The researchers found a correlation of .89 between the amount of contingent instruction the mother used and the child's performance; the more contingent the instructions, the more the child learned. Wood replicated these results in a second study (although the correlations were somewhat lower (.65)) that examined not only the tutorial interaction between a mother and her child, but also between the mothers and children they did not know. He concluded that the instructional strategy was a more powerful determinant of effective instruction than the relationship between the mother and child. A final experiment was conducted (Wood, Wood, & Middleton, 1978) to determine if the contingency rule was a crucial part of the tutoring process. A single experimenter-tutor was trained to use four different tutorial strategies. The investigators found that children who were taught contingently learned much more than children who were taught by any of the other methods.

In this study, an adaptation of the experimental paradigm developed by Wood, et al. (1978) was used to test two hypotheses:

1) That the effectiveness of contingent instruction generalizes beyond the task (three-dimensional puzzle) and the population (3, 4, & 5 year olds) in the Wood et al. studies.

2) That it is possible to teach these tutoring techniques. (i.e., The instructional effectiveness is not just a function of the characteristics of the individual tutor.)

METHOD

Subjects.

College students were recruited from sections of an Elementary Algebra course (equivalent to first year high school algebra) at a large state university. The students

enrolled in the course represented a diverse cross-section of the university on the dimensions of age, gender, race and ethnic background. The 56 subjects who participated in the study were mainly freshmen and sophomores, ranging in age from 18 to 31 years ($\bar{X}$ = 21.5) The one distinguishing characteristic of all students enrolled in this course was that they had a weak background in mathematics, which was the reason for selecting this population. It was unlikely that students who were currently taking elementary algebra would have ever taken any physics, the topic of the experimental tutoring session. In addition, students enrolled in "remedial" courses are typically representative of the target populations that tutoring programs on college campuses are designed to serve. Volunteers for whom SAT scores were available were chosen to be subjects. Subjects were randomly assigned to experimental conditions, stratifying on race, gender, and aptitude (combined SAT scores) to guarantee representativeness within each condition, as described in Figure 1. Subjects were paid $10.00 for their participation.

Insert Figure 1 here

Design and Procedure.

Two experienced tutors, were trained to tutor college undergraduates in a basic optics lesson on the nature of lenses (adapted from Collins, 1987). Each tutor taught the same optics lesson using three different instructional strategies in a 2 (tutor)x 3 (instructional strategy) factorial design, with eight subjects per cell.

In each condition the instruction was heavily scripted to ensure that the content taught would be consistent across conditions. What varied was how the material was taught. The three instructional conditions were modeled after three of the conditions developed by Wood, et al. (1978), varying the amount of control exerted by the tutor. The "Discovery" condition required the student to assume almost full responsibility for learning. The tutor provided only

positive or negative feedback, or high level verbal instructions, either encouragement--to"prompt activity without providing any specific indications of the goals to be achieved," or to establish parameters verbally, by directing attention. This condition was similar to traditional "discovery" techniques that advocate learning by doing, through loosely-guided exploration.

In the "contingent" condition, control over the learning was shared by the tutor and the student. The questioning strategy adopted by the tutor provided constant interaction between the tutor and the student. Each tutorial intervention was contingent upon the previous response of the student. When the student made an error, or failed to comprehend, the tutor adapted by increasing control, providing more direction to help the student understand. When the student was successful, demonstrating comprehension, the tutor relinquished some control. The major idea was to encourage the student to do as much as possible on his or her own.

In the "Lecture" condition the tutor assumed complete control over the instruction. The tutor "taught" the student by explaining how to do all aspects of the task. This condition is similar to the type of lecture instruction typically used in schools. Phenomena were demonstrated for the student and the processes involved were explained.

A fourth group of eight students was asked to learn the same lesson from a written text. Again, the content of the lesson was identical to the three tutoring conditions, and utilized the same diagrams, although students in the baseline condition did not have the benefit of the materials to manipulate and explore. This condition served as a baseline measure, against which the effectiveness of the three tutoring strategies could be compared.

Subjects received the same set of instructions and were given the following set of learning objectives as an "advanced organizer" before the instruction began.

1. To be able to identify any concave or convex lens.
2. To be able to define and identify the focal point and focal length of any convex lens.

3. To be able to show the path of light graphically as it passes through different parts of any convex lens.
4. To understand the basic relationship between focal length, lens curvature, and magnification.

The Task

Optics was chosen because it was expected to be a novel task for all students in the population. Optics is not part of most high school science curricula in either biology or chemistry. The topic is first encountered in introductory physics courses, and it is highly improbable that students enrolled in first year algebra would have had any physics training. A brief pretest was administered to test this assumption and control for prior knowledge. More importantly, however, optics was chosen because it is a structured domain that can be objectively evaluated. There was also an existing literature of task analyses and lesson prescriptions for the topic of how lenses work (Goldberg & McDermott, 1987; Reigeluth, 1987). Moreover, to be a fair test of discovery methods it had to be a task whose solutions were empirically confirmable, and where the correct response could be logically deduced from the demonstrations, diagrams, and ideas presented to the students (Cronbach, 1966).

The lesson on lenses was adapted from Collins (1987), incorporating material from numerous other sources (Goldberg & McDermott, 1987; Hewitt, 1981; Miller, 1987; Van Name & Flory, 1974; Wilson, 1978). A progressive sequence of presentations and actvities were developed using a set of four diagrams, a variety of concave and convex lenses, and a simple home-made optical bench, which allowed students to observe and measure the focal length of the convex lenses.

Following the session each subject took a posttest to determine how much the student learned from the lesson. The students were then asked to solve two related, but unfamiliar

problems. During the problem-solving session the tutor asked the students to explain their reasoning. These questions were designed to provide some insight into how the students had organized the information, and the gaps that might exist in the student's knowledge structures (both declarative and procedural). If the students could not solve the problem on their own, the tutor tutored the student until a solution was reached. At the conclusion of the session subjects were asked to fill out a tutor evaluation. The entire experimental session lasted approximately one and one-half hours. All sessions were videotaped.

Coding the videotapes

Two independent coders, who were blind to the hypotheses provided a reliability check on the fidelity of the manipulations of the independent variables. The coder's scoring of the videotapes was used to determine whether each tutor performed consistently according to the prescribed instructional strategy. Each coder scored all of the tutoring sessions.

A six-category hierarchical coding scheme was developed (adapted from Wood, et al. 1978) to analyze the videotapes of the tutoring sessions. Each tutor intervention was classified in terms of the degree of control exerted by the tutor over the student's next move. The coding system, presented in Figure 2, constitutes a plausible ordinal scale of "amount of control" applied by the tutor in the instructional interaction. Reading down the list, the categories of instruction become more specific, informative, and controlling.

Insert Figure 2 here.

When the student works alone, with only the tutor's encouragement, such as 'you're doing great,' (Level 0), the tutor does not provide any specific guidance as to what to do or how to do it. The student is in full control of the learning activity. If the tutor asks, for example, 'Why are the focal lengths of these lenses different?', (Level 1) the tutor is pointing to a critical

feature, but is still asking the student to figure it out for herself. By asking the student, 'Which of these lenses is the most curved? (Level 2), the tutor becomes more controlling by directing the student's attention to specific features. If the student needs additional assistance, the tutor might break down the problem into smaller, more manageable steps (Level 3). For example, 'If we compare these two small sections of the lens, (pointing to specific parts of a diagram), which part of the lens is the most curved? And which part bends the light the most?' Often students need even more direction. The tutor might explain the objectives or organization of the task to direct the student through the task (Level 4). An example of a Level 4 intervention might be, 'In this diagram the lens is much more curved. I want you to complete the picture by drawing the way the light rays will travel through this lens.' Finally, if the tutor gives an explanation of the process, or provides an answer (e.g., 'The more curved the lens is, the more it will bend the light.'), the tutor has taken over full responsibility and control over that phase of the task.

In addition, to capture the transactional nature of the tutorial dialogue, the coders scored the students' responses to each tutor intervention. Students' responses were coded to indicate whether the student understood (+), or failed to comprehend (-). In many cases it was impossible to judge whether the student really understood the tutor's point. In these instances the student's response was coded (?), indicating that it was ambiguous. In the case where there was no response from the student, the cell was left blank.

The completed coding sheets then, provide a detailed description of each tutoring session, and supply the necessary data to evaluate the consistency with which each tutor carried out the specified instructional strategy. For example, in the lecture condition the tutor was supposed to maintain complete control over the lesson by explaining to the student all about how lenses work. Therefore, the code sheet from a lecture session would contain only Level 4 and Level 5 codes, (explanations of how the task is organized, and how it works). In contrast, in the discovery condition, the tutor was to provide only Level 0 and Level 1 instructions, basically

only providing feedback on the student's work, and suggesting strategies that students might use to figure out for themselves how lenses work.

Evaluating contingent instruction is more complex. In the contingent condition the tutor's goal was to provide instruction that was contingent upon the momentary successes and failures of the student. Contingency is defined according to two rules. If, following an intervention, the student does not succeed, the tutor should provide more help or control. For example, if a leading question (level 2) fails to elicit a correct response, the appropriate next move would be to break down, or simplify the problem (Level 3). Alternatively, if the student responds correctly to the question, the tutor should decrease control, by asking the student to explain how or why a phenomena occurred (Level 1), or offer only general encouragement (Level 0), until a problem is encountered that the student cannot solve. By following these two rules, the tutor attempts to confront the student with problems at the appropriate level of difficulty. Increasing help when the student is struggling provides the student with a more manageable problem. Stepping back when the student demonstrates understanding, implicitly increases the demands placed upon the student.

Wood, Wood, & Middleton (1978) argue for a strict interpretation of the contingency rule when evaluating tutorial dialogues. When the tutor stayed with the same level of intervention after a success or failure, instead of immediately increasing or decreasing control, these researchers counted the 'transgression' as an error. Wood, et al. found that there was a general tendency for tutors to give the child more opportunities to note or correct his own mistakes without an increase in instructor control. Although this seems a natural and familiar ploy, earlier work (Wood & Middleton, 1975) suggested that such transgressions were likely to diminish the effectiveness of the instruction. While this hypothesis is worthy of further investigation, in the present study, given the complexity of the subject matter taught, a more liberal scoring strategy was adopted. Tutor interventions were only counted as errors if the

tutor increased control after a success, or decreased control after a failure. It was however, considered an error if the tutor continued to maintain control by providing instruction in very small steps for an extended period of time without providing the student with opportunities to assume more control over the learning process. In addition, if the tutor provided an answer or explanation (Level 5) without gradually increasing control, and breaking the problem down to small manageable units first, it was considered an error. Inter-rater reliability for the total sample of 4,941 interventions was high (92.9% agreement).

RESULTS AND DISCUSSION

The results were analyzed from two perspectives. The first examined the main effects of tutor and instructional strategy on students' learning. The second point of view explores individual differences in response to the different instructional methods.

As expected, prior knowledge of the subject matter was limited and idiosyncratically distributed. The modal score was 0, indicating that the topic was relatively novel to most students. Pretest scores were generally low, $\bar{X} = 5.07$, and uncorrelated with any of the aptitude measures or the posttest scores. Overall, the correlation between pretest and posttest scores was very low (.16).

Although based on a small sample, the relative effects of the three methods of instruction on student achievement were consistent with prior research findings. Posttest means and standard deviations, overall and by condition and tutor, are reported in Table 1. The mean posttest score for the entire sample was 24.02. While treatment differences on the posttest were small, group averages were consistently lowest in the discovery condition ($\bar{X} = 22.56$). On the average, students in the contingent condition obtained somewhat higher posttest scores than students in the lecture condition, but averages can be misleading, and obscure real differences in students' performance. Posttest scores were consistent across tutors in the lecture and discovery conditions, but inconsistent in the contingent treatment. Students

tutored by Tutor B in the contingent condition earned the highest posttest scores ($\overline{X}$ = 27.38), while students tutored by Tutor A scored slightly below the overall mean ($\overline{X}$ = 22.13). Differences in mean outcomes describe only general differences among treatments, but ignore individual differences in response to those treatments. To address both issues, mean outcomes must be considered in conjunction with regression slopes.

Insert Table 1 here

The results of the regression analysis on posttest are presented in Table 2. An R^2 of .59 was obtained for the full model. Aptitude main effects, specifically SAT scores, accounted for the largest portion of explained variance (36%), indicating that high ability students consistently achieved higher scores than did low ability students. None of the treatment main effects were significant, nor was the main effect of tutor significant. Differences in the mean outcomes of the four conditions were small, as were the differences between tutors A and B.

Insert Table 2 here

There were however, significant first-order Aptitude Treatment Interaction (ATI, Cronbach & Snow, 1977) and Treatment x Tutor interactions. Classic Aptitude Treatment Interaction results were observed (See Figure 3). The steepest regression slope was found in the discovery condition, where low ability students were particularly handicapped by this method of instruction. The discovery slope is strongly influenced by two outliers, two low ability students who experienced great difficulty with this mode of instruction. It should be noted, however, that the performance of these two students was consistent with the trend observed with this treatment: the lower the ability, the weaker the performance. This

observation is also consistent with previous research on discovery learning (Cronbach, 1966; Cronbach & Snow, 1977). Instructional treatments that place heavy information processing burdens on the learner, as in the discovery condition, produce relatively steep regression slopes of learning outcomes onto ability. Able students do well and less able students do poorly. When the instruction is fully elaborated, as in the lecture condition, the regression slope was less steep. And more adaptive teaching methods, such as the contingent treatment, reduced the regression slope even further, narrowing the gap between low and high ability students.

Insert Figure 3 here

The Treatment x Tutor interaction was also significant, accounting for 6.1% of the variance. The interaction becomes obvious when the regression of SAT on posttest is calculated separately by tutor. Figure 4 A and B graphically illustrate these results. Although each regression line represents only eight students in each condition, marked differences appear. For Tutor A the lecture strategy produced the highest posttest scores ($\overline{X}$ = 24.25), while the contingent treatment produced somewhat lower learning outcomes ($\overline{X}$ = 22.13). In contrast, for Tutor B the order of achievement was reversed ($\overline{X}$ = 27.75 in contingent, versus $\overline{X}$ = 24.63 in the lecture condition). The contingent strategy yielded higher outcomes, and a flatter regression slope. Low ability students seemed to profit most from this instructional approach, but it was the most effective treatment among high ability students as well. Students tutored by Tutor B using the lecture technique obtained intermediate results. Comparable steep slopes were observed for both tutors in the discovery condition.

Insert Figure 4 here

Consistency of Tutor Interventions

To understand the different outcomes obtained by each tutor, the videotapes of the tutoring sessions were analyzed. Inspections of the videotapes revealed that the planned instructional treatment was not always the treatment delivered. Both tutors administered the lecture strategy with a high level of consistency. Tutor A's interventions complied with the treatment constraints 96% of the time, while 95% of Tutor B's interventions were consistent. However, in the contingent condition, where the largest between-tutor differences were observed, Tutor A was successful in complying with the contingency rule only 65% of the time, while 97.4% of Tutor B's interventions were consistent. Proportions may not give an accurate picture, as each intervention does not have the same instructional impact. Leading questions that provide stronger hints than needed still require the student to do some of the intellectual work; providing answers does not. Perhaps a more revealing statistic is the number of inappropriate explanations given by the tutors. Tutor A explained ideas or provided answers an average of seven times per session (Range: 3-12), where he should have broken down the problem gradually into more manageable steps that would allow the student to come to the correct solution. Tutor B, on the other hand, gave an average of only two inappropriate explanations (Range: 0-4) each session. A subsequent analysis showed a postitive rank-order correlation between each tutor's contingency scores and students' posttest performances (Tutor A: $r = .93$, $N=8$, $p < .05$; Tutor B: $r = .96$, $N=8$, $p < .05$). For both tutors, the more contingent the instruction, the better the students' performance. These results demonstrate support for the effectiveness of contingent instruction, when the strategy is correctly applied.

The greatest number of errors actually occurred in the discovery condition, where both tutors found it difficult to refrain from providing more help than was specified by the instructional method. Tutor A did more teaching than was "allowed" 41.7% of the time. Tutor B was somewhat more consistent with the treatment constraints, but still provided more help than

was called for 27% of the time. Strict compliance with the principles of discovery learning might have produced even greater decrements in performance. When each discovery tutoring session was rank ordered and compared with outcomes, significant rank-order correlations were found (Tutor A: $r = -.94$, $N=8$, $p < .05$; Tutor B: $r = -.75$, $N=8$, $p < .05$). Thus, the less successful the tutors were in carrying out the prescribed treatment (i.e., the more help given), the better the students' performance.

Figure 5 shows graphically that real differences did occur between the two tutors in the contingent condition. For the entire sample the mean differences between conditions were small, but for Tutor B (the tutor most successful in following the rules specified in each condition), there were significant outcome differences between the three instructional strategies, ($F(2, 21) = 3.54$, $p < .05$). The "Contingent" group scored .65 standard deviations above the mean, the "Discovery" group scored a half a standard deviation below the mean (-.52), while the "Lecture" group performed approximately at the mean (.14). There were no significant differences between conditions among students tutored by Tutor A, perhaps because there was less differentiation between the strategies as administered by this tutor.

Insert Figure 5 here

Performance on Novel Problems

Scores for the problem solving sessions were inconsistent and provided little insight into the depth or quality of students' understanding of optics. One possible explanation may be that the "novel" problems were not novel enough to evaluate students' understanding of the processes involved. A "novelty" index is necessarily idiosyncratic and relative to students' knowledge of the topic. Although the problems presented to the students were new (i.e., they were not taught during the tutoring session), they were about familiar, everyday phenomena,

and this familiarity seems to have complicated the issue. The fact that the problems were practical problems had both positive and negative consequences. The practicality provided motivation and interest. Students commented that it was interesting to see how the theoretical principles applied to common situations from their own experience. A drawback of using practical problems is that some students had "common" knowledge of how telescopes or the eye worked, and were able to solve the problems without applying the principles taught. Retrospective interviews revealed that some students were able to solve the problem because they "already knew how the eye worked." However, when asked to explain how the shape of the lens would effect the path of light, they could not explain the process.

There was also inconsistent administration of the two problems by one of the tutors. Both coders commented that Tutor A's questions were often vague and poorly worded.

Tutor Evaluations

Students' ratings of both tutors were extremely positive, on a scale of 1-5, where 5 is the most favorable, the mean score for Tutor A was 4.25 (SD = .49), and the mean for Tutor B was 4.56 (SD = .44) The small difference between tutors was not significant, ($F(1.46) = 1.24$, $p > .6$). Tutor differences were found on only two items. Students working with Tutor B felt more comfortable working ($\bar{X} = 4.54$), than did students working with Tutor A ($\bar{X} = 4.08$) ($F(1,46) = 3.59$, $p < .003$). In addition, students tutored by Tutor B gave a stronger endorsement of individual tutoring as an effective way to teach (Tutor B $\bar{X} = 4.88$, Tutor A $\bar{X} = 4.40$, $F(1,47) = 5.84$, $p < .001$). Comments made in open-ended questions correspond with these figures. Six of Tutor B's students commented on aspects of her tutoring that made them feel comfortable during the lesson, particularly that she was non-judgmental. "She made me feel good when I got it right," and she "didn't berate me for wrong answers, rather encouraged me to understand the right answer," and she "corrected 'mistakes' without making me feel stupid." These remarks were not indicative of any one instructional treatment; similar comments came

from students in each of the three conditions. Thus, this difference is most likely a function of the tutor's personal style rather than her implementation of a particular instructional strategy.

Other student comments praised both tutors for being knowledgeable and patient, for the clarity of the instruction, and for "making me think". Students ranked patience as the quality most important in a tutor, followed by communication skills, and knowledge of the subject matter.

Summary of Results

In addition to the inconsistent implementation of the instructional treatments, conclusions drawn from this study are limited by the small sample size and short duration of the experiment. Each tutor tutored only eight students using each of the instructional methods, and the number of relatively high or low ability subjects within each treatment was smaller still. Larger samples are needed to test the effects of instruction on the relation of aptitude and learning outcomes. In particular, additional research is needed to confirm the adaptive power of contingent instruction across the ability spectrum. The results of this experiment are, however, noteworthy in that it is unusual to find such definitive interactions in an experiment with such a short period of instruction (25-50 minutes), and such a small sample.

The results of the present experiment are also consistent with earlier investigations of contingent instruction (e.g., Wood, 1980; Wood, Bruner & Ross, 1976; Wood, Wood, & Middleton, 1978). When the strategy is correctly applied, it is the most effective tutoring strategy. These results are also consistent with previous ATI research. Adaptive instructional methods, such as contingent instruction, decreased the relationship between aptitude and achievement, and produced the greatest learning gains. Discovery learning, on the other hand, was the least effective strategy for teaching how lenses work, particularly among less able students. Lecture techniques produced intermediate results.

Contingent instruction appears to be effective because it provides the optimal level of

structure and assistance for each individual student. This tutoring technique, however, proved to be quite difficult to learn and carry out consistently.

Implications for Tutor Training

The inconsistent administration of the contingent strategy underscores the subtlety of this tutoring technique, which requires tutors to be sensitive to the needs of individual students. It is, of course, too soon to offer prescriptions for tutor training based on the results of a single study. If, however, these findings hold up in future investigations, then tutor training programs will be needed to teach tutors to use contingent strategies. Although the attempts to train research assistants to do so in this study were not entirely successful, much has been learned to advise future efforts in tutor training. With hindsight, it is clear that the training efforts in this study were both hasty and superficial. It turned out that the contingent strategy was much more sensitive and subtle than even the author, who designed the instruction and trained the tutors, had originally understood.

The abstract notion of contingent responsiveness was often difficult to put into words. For tutors new to this approach, just phrasing their instruction in the form of questions took a great deal of practice. In working with the tutors efforts were made to provide examples and possible alternative questions to demonstrate how to adjust the level of their questions in response to students' answers. Although the tutors seemed to understand the explanations, formulating questions "on-line" in the flow of tutorial interactions was a particularly difficult skill to master, largely because it was impossible to predict the kinds of questions and comments that students would come up with, or when they would respond with only a blank stare. The author was present during all of Tutor A's pilot sessions, and after observing, critiquing, and discussing 12 cases, felt confident that he had achieved an acceptable level of consistency in applying the strategy. Although observations of his interactions with the experimental subjects revealed "errors" in the contingent condition, there were obvious differences in his

characterization of the three strategies, and it was thought that the differences were profound enough to produce differences in students' achievement.

Most of the errors made by both tutors in the contingent condition were failures to relinquish control when the student demonstrated understanding. This was particularly true of Tutor A, who understandably adopted a strategy that was safe--less adaptive, but less uncertain. By breaking down the instruction into very small steps, he maintained a high level of control and decreased the need to improvise in response to unexpected and erratic student replies. To some students these small steps rendered the answer "obvious" and did not challenge them to struggle with the ideas. Without a level of uncertainty in the problems posed in the lesson, students failed to develop a deep understanding of the processes involved.

An important point is that Tutor A's lack of sensitivity to the responses of individual students was not recognized until the videotapes were analyzed. Only in the process of analyzing the tapes did the lack of fit between his interventions and the student's level of understanding become apparent. This underscores the need for examining the practices of individual tutors in a variety of situations. Tutors, themselves, might benefit from systematic observations of the effects that their teaching has on students. A training program that incorporates 'self-observation' studies where tutors code videotapes of their own tutoring sessions would not only help tutors recognize tutoring strategies that seem most productive, but would also help them to see the tutoring process from the students' point of view.

A successful training program of this sort has been developed by Wood, and his associates (1986) to train teachers of the deaf. In the year-long course students engage in a 'longitudinal' approach to evaluating their own tutoring strategies. By evaluating videotapes throughout the year they are able to assess changes and improvements in in their tutoring technique. Trainees are also shown videotapes of other tutors who demonstrate both effective and ineffective styles. The experience and suggestions of both peers and supervisory staff provide a

supportive context for developing effective tutoring skills.

Implications for the Development of Intelligent Tutoring Systems

It is fairly clear the ITS cannot replace human tutors in the foreseeable future. While efforts at designing artifically intelligent instructional systems will benefit from careful analyses of how concepts emerge in the teacher-student interaction, current computer systems are actually quite far from being able to perform the feats of sensitive interpretation performed by human teachers. A recommendation that follows from this research is that the design of intelligent tutoring systems should not attempt to replace the teacher, but rather ITS should provide ways of augmenting and extending the power of teachers. With the support of the computer, the teacher can assume the role of an adaptive tutor, mediating between the student and the machine. This set up would enable a more flexible strategy that would give the student a more active role in determining the goals and the sequence of instruction, and also allow the teacher to be more responsive to varying needs of individual students.

It should be possible to adapt contingent teaching strategies to any instructional setting, presumably with positive results, as long as its basic principles are maintained. First, students must be presented with problems that are within their zone of proximal development. Students must be able to recognize and understand the goals of the task before they can begin to achieve them. Second, by adopting a problem solving approach to instruction, the teacher can adjust the level of intervention to the needs of the students-- providing sufficient support to allow the students to do as much as possible on their own. And finally, successful instruction requires a progressive relaxation of teacher control over the problem solving process. Each success will help students build confidence in their own abilities, producing competent and independent learners.

It is rarely the case that one instructional strategy is best for all people. Numerous ATI studies have demonstrated that students with different aptitude profiles benefit differentially

from different teaching methods. Specifically, a consistent ATI finding is that instructional treatments that place heavy information processing demands on the learner are better suited to more able students, while less able students tend to benefit when the instruction is more structured and fully elaborated. The results of the present study are consistent with these earlier findings. Discovery learning, which requires the student to assume full responsibility for learning, was particularly debilitating for less able students, widening the gap between low- and high-achieving students. In the lecture treatment, the tutor assumed full control of the lesson, explaining all aspects of the task, thereby requiring the student to do little intellectual work. Less able students benefited from the more complete instruction, and as a result, the slope of the regression of learning outcome on ability was reduced. The contingent condition used a more adaptive approach, the degree of tutor control varied with the students' ability to handle the material. This strategy was particularly helpful for less able students. Moreover, when correctly administered, contingent instruction produced higher posttest scores among students of all ability levels. Perhaps the strength of this technique is that it is essentially adaptive.

Implications for Future Research

Only long-term studies with substantially large samples will be able to confirm or refute the conclusions of this study. Additional research is necessary to address the generalizability of these findings to other student populations and other subject domains. If developing the learning skills of students is the ultimate goal of education, then a better understanding of the effects of instructional practices on students' reasoning skills is needed. To do this, researchers will need to examine both immediate and delayed outcomes to determine the power of various instructional methods to promote long-term retention and transfer.

By experimentally evaluating the efficacy of tutorial strategies, this study has provided some additional insights into the ingredients that contribute to effective instruction. However, many questions remain unanswered. In particular, little is known about the influence

of interpersonal factors in the tutoring process. How does the relationship between the tutor and the student affect the student's motivation and effort? How do tutors develop trust and rapport with their students? A first step toward understanding these processes are studies that examine how understanding develops in the course of tutor-student interactions. Descriptions of what human tutors do is an important step in this direction, but it is only a first step. It is not only necessary to study what human tutors do, the research must go one step further and construct experiments to evaluate the effectiveness of these practices in helping students learn.

REFERENCES

Bloom, B. S. (1984). The 2-Sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher, 13, 4-16.

Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. J. Jarvell, & W. J. M. Levelt (Eds.), The child's conception of language. (pp. 241-256). New York: Springer.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal, 19, 237-248.

Collins, A. (1987). A sample dialogue based on a theory of inquiry teaching. In C. M. Reigeluth (Ed.)., Instructional theories in action: Lessons illustrating selected theories and models. (pp.181-199) Hillsdale, N.J.: Erlbaum.

Collins, A., & Stevens, A. L. (1982). Goals and strategies of inquiry teachers. In R. Glaser (Ed.), Advances in instructional psychology, (Vol. 2) Hillsdale, N.J.: Lawrence Erlbaum.

Cronbach, L. J. (1966). The logic of experiments on discovery. In L.S. Shulman & E. Keislar (Eds.) Learning by discovery: A critical appraisal. (pp. 76-92). Chicago: Rand McNally.

Cronbach, L. J., & Snow, R. E. (1977). Aptitudes and instructional methods. New York: Irvington Publishers.

Goldberg, F. M., & McDermott, L. C. (1987). An investigation of student understanding of the real image formed by a converging lens or concave mirror. American Journal of Physics, 55, 108-119.

Hewitt, P. G., (1981). Conceptual physics: A new introduction to your environment. (4th ed.). Boston: Little, Brown and Company.

Lepper, M. R., & Chabay, R. W. (1988). Socializing the intelligent tutor: Bringing empathy to

computer tutors. In H. Mandl & A. Lesgold (Eds.), Learning issues for intelligent tutoring systems. (pp. 242-257). New York: Springer.

Littman, D., Pinto, J., & Soloway, E. (1987). The knowledge required for tutorial planning: An empirical analysis. Paper presented at the Third Annual Conference of Artificial Intelligence and Education, Pittsburgh, PA.

McKendree, J., Reiser, B. J., & Anderson, J. R. (1984). Tutorial goals and strategies in the instruction of programming skills. Paper presented at the Sixth Annual Conference of the Cognitive Science Society, Boulder, CO.

Miller, B. (1987). Bob Miller's image walk. Exploratorium Quarterly, 11, 2-12.

Van Name, F. W., Jr., & Flory, D. (1974). Elementary Physics (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Wilson, R. (1978). How the body works. New York: Larousse and Co., Inc.

Wood, D. J. (1980). Teaching the young child: Some relationships between social interaction, language, and thought. In D. R. Olson (Ed.), The social foundation of language and thought. (pp. 280-296). New York: W. W. Norton.

Wood, D. J., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. Journal of Child Psychology and Psychiatry, 17, 89-100.

Wood, D. J., & Middleton, D. J. (1975). A study of assisted problem solving. British journal of psychology, 66, 181-191.

Wood, D. J., Wood, H., Griffiths, A., & Howarth, I. (1986). Teaching and talking with deaf children. New York: John Wiley & Sons.

Wood, D. J., Wood, H., & Middleton, D. (1978). An experimental evaluation of four face-to-face teaching strategies. International Journal of Behavioral Development, 1, 131-147.

Instructional Style	Tutor A	Tutor B
Lecture	4 MALES 4 FEMALES 3 WHITE 5 NON-WHITE SAT : $\bar{X}$ = 831 SD = 178 N=(8)	3 MALES 5 FEMALES 4 WHITE 4 NON-WHITE SAT : $\bar{X}$ = 840 SD = 168 N=(8)
Contingent	3 MALES 5 FEMALES 3 WHITE 5 NON-WHITE SAT : $\bar{X}$ = 851 SD = 148 N=(8)	3 MALES 5 FEMALES 4 WHITE 4 NON-WHITE SAT : $\bar{X}$ = 853.5 SD = 176 N=(8)
Discovery	4 MALES 4 FEMALES 5 WHITE 3 NON-WHITE SAT : $\bar{X}$ = 842.5 SD = 179 N=(8)	4 MALES 4 FEMALES 4 WHITE 4 NON-WHITE SAT : $\bar{X}$ = 838.8 SD = 157 N=(8)

N = 48

Baseline	3 MALES 5 FEMALES 5 WHITE 3 NON-WHITE SAT : $\bar{X}$ = 837.5 SD = 103.5 N=(8)

Figure 1. Subject distribution, stratified by gender, ethnic background, and combined SAT scores.

Figure 2. CODING SCHEME DESCRIBING LEVELS OF TUTOR INTERVENTIONS

LEVEL 0 - Student works alone; Tutor provides encouragement and reinforcement.

LEVEL 1 - Tutor points out critical features to use as checks

--suggesting strategies

--"why" questions; ask student to explain process

LEVEL 2 - Provide direction, guide student to next step

--leading questions; use questions to direct student what to do next.

LEVEL 3 - Simplify task so student can focus on manageable unit.

--break down problem into small steps, or isolate a small part

LEVEL 4 - Explain objectives and organization of task

--explain what to do

--explain diagram

LEVEL 5 - Give explanation and answer.

--tell student how it works.

TABLE 1. POSTTEST MEANS AND STANDARD DEVIATIONS

N = 8 Subjects per cell

CONDITION

	LECTURE		CONTINGENT		DISCOVERY		BASELINE		OVERALL	
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD
TUTOR A	24.25	4.43	22.13	4.45	22.75	5.83	- - -	- - -	23.04	4.81
TUTOR B	24.63	2.93	27.38	2.56	22.13	6.34	- - -	- - -	24.92	4.54
OVERALL	24.44	3.63	24.94	4.52	22.56	5.79	24.25	4.74	24.02	4.70

Table 2. Summary of regression of Posttest.

VARIABLE	d.f.	*B*	% VARIANCE ACCOUNTED FOR	F-RATIO
FULL MODEL	12	--	59	7.10*
APTITUDE MAIN EFFECTS	2	--	39	16.71*
PRETEST	1	.109	3	1.94
SAT (COMBINED)	1	.022	36	31.23*
TREATMENT MAIN EFFECTS	3	--	3.4	1.79
T1 (LECTURE)	1	.326	<1	.68
T2 (CONTINGENT)	1	11.75	1.4	1.24
T3 (DISCOVERY)	1	-1.74	1.3	1.13
TUTOR MAIN EFFECTS	1	--	1.9	1.69
TUTOR	1	.2	1.9	1.69
FIRST ORDER ATI	3	--	8.4	5.63*
SAT X T1	1	-.013	<1	1.06
SAT X T2	1	-.018	7.7	8.51*
SAT X T3	1	-.003	0	.05
FIRST ORDER TREATMENT X TUTOR INTERACTION	3	--	6.1	3.35*
T1 X TUTOR	1	.199	1.2	1.06
T2 X TUTOR	1	-2.83	4.9	4.89*
T3 X TUTOR	1	.00	0	.00
RESIDUAL	46	--	41	--

*p< .05

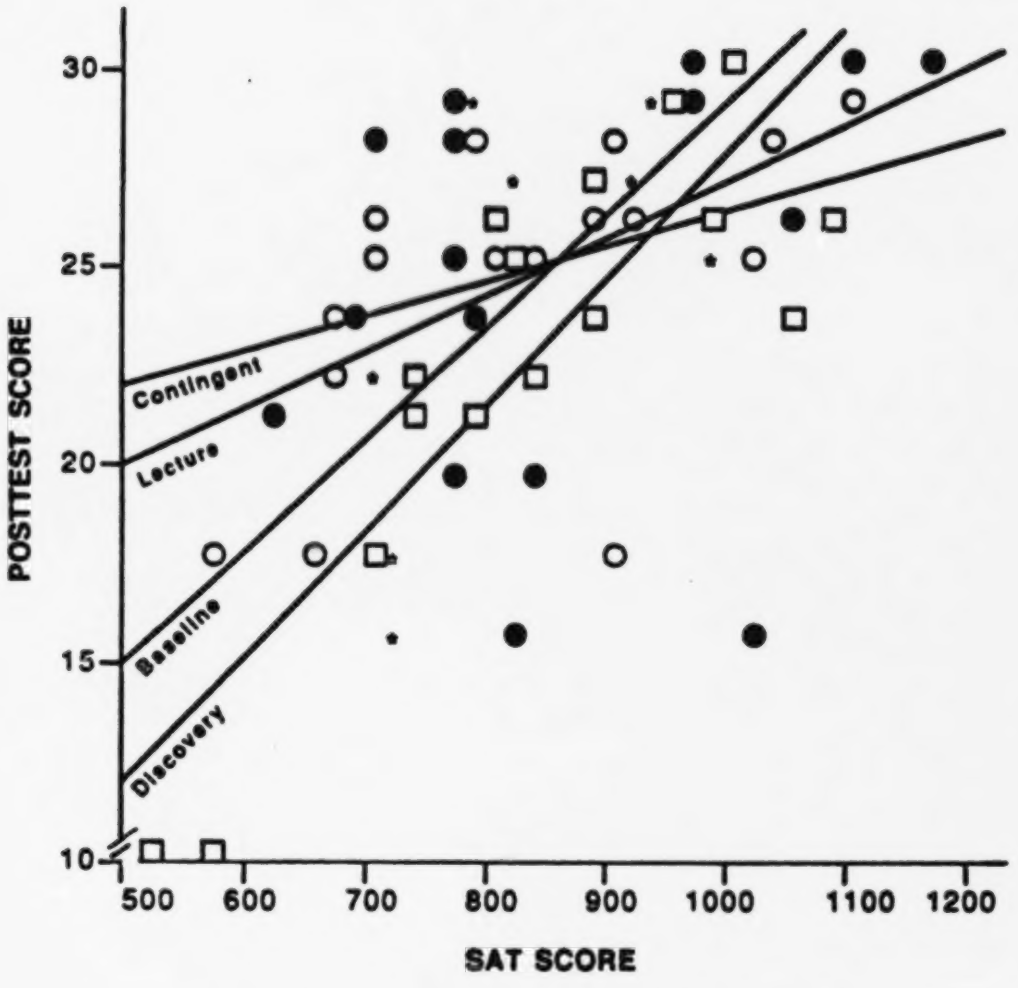

Figure 3. Relation of SAT (combined) to Posttest By Condition

○ Lecture ● Contingent □ Discovery • Baseline

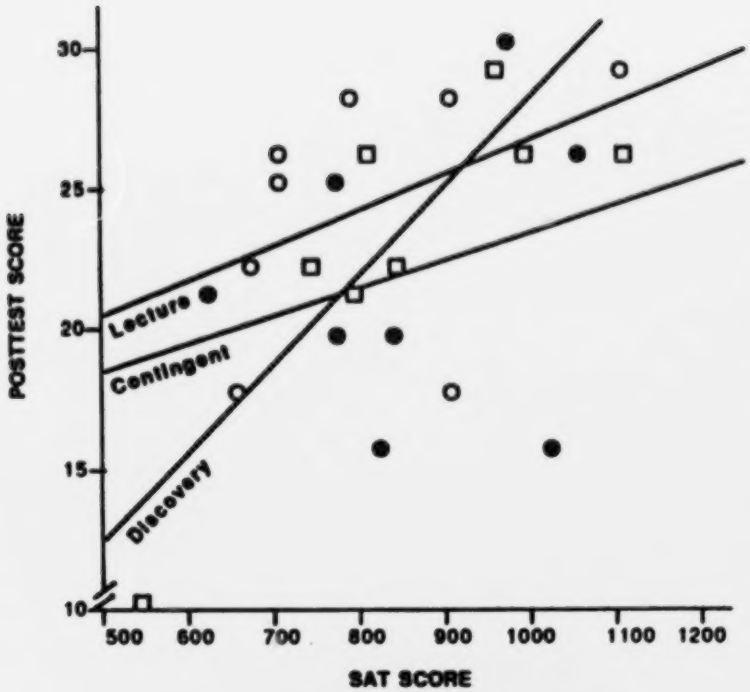

Figure 4. A Relation of SAT (combined) to Posttest By Condition For Tutor A.

○Lecture ●Contingent □ Discovery

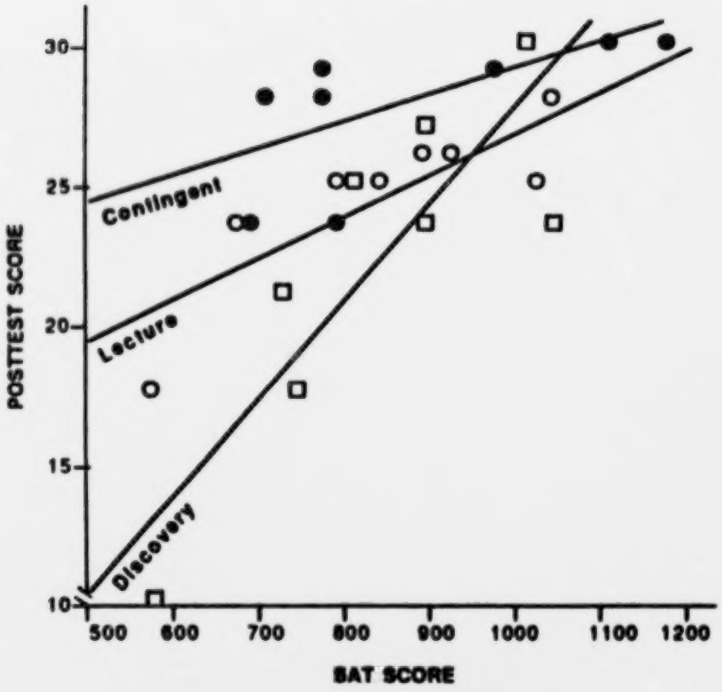

Figure 4 B. Relation of SAT (combined) to Posttest By Condition For Tutor B

○Lecture ●Contingent □ Discovery

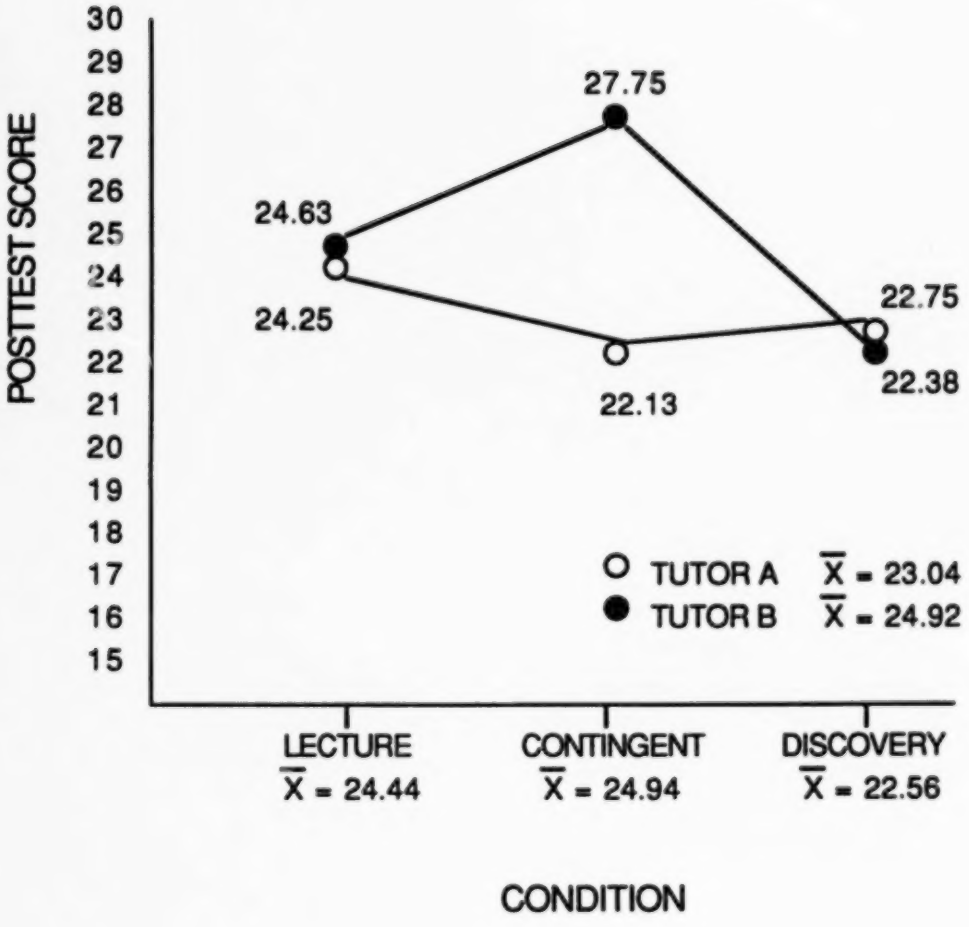

FIGURE 5. POSTTEST MEANS BY TUTOR AND CONDITION